Advisory committee briefing document: DHEA NDA 21-239

TABLE OF CONTENTS

REVIEW

1. CLINICAL DHEA EXPOSURE IN THIS NDA

The clinical exposure (i.e. excluding brief exposure for PK purposes) in this NDA consists of patients in four trials:

- **GL94-01 – 7-9 month RCT of 191 females with mild-mod. lupus**
- **GL95-02 – 1-year RCT in 381 females with mild-mod. lupus**
- **GL95-01 – 1-year open trial, for patients completing GL94-01 or GL95-02 (extended a 2nd year in one center)**
- **GL97-01 – 1-year RCT, 1-year open study in 40 males (ongoing / 28 enrolled to date)**

Patients in either of the two main efficacy trials in this submission, GL94-01 and GL95-02, were given the option of continuing DHEA in an open extension (GL95-01) at 200mg/d. Table 1A shows the flow of patients from trials GL94-01 and GL95-02 to trial GL95-01. The open extension had a duration of 12 months, except for one center (Johns Hopkins), where some patients were continued a second year. Table 1B enumerates overall exposure at 200mg/day.

Table 1A: SOURCE OF EXPOSURE
number of patients

	# entered RCT	# completed RCT	---open: Trial GL95-01 - 200mg/d----- # entered	0-6mo	6-12mo	12-18mo
RCT#1-GL94-01						
DHEA200mg	**64**	**47**	**43**	**8**	**30**	**5**
DHEA100mg	**63**	**46**	**44**	**6**	**34**	**4**
PLC	**64**	**49**	**46**	**8**	**35**	**3**
RCT#2-GL95-02						
DHEA200mg	**189**	**124**	**105**	**19**	**80**	**6**
PLC	**192**	**142**	**133**	**34**	**94**	**5**

TABLE 1B: TOTAL 200mg/d EXPOSURE

Time period	Number of patients
>6 months	**214**
>12 months	**118**
>24 months	**11**

2. Clinical Background

Systemic lupus erythematosis is a multisystem disease commonly affecting the skin, joints, and kidney, and it is often accompanied by constitutional symptoms such as malaise and fatigue. It is much more common in females (approximately 8 to 1), with a tendency to worsen at menarche and lessen at menopause. Historically, renal lupus has accounted for major morbidity and mortality, and has been the focus of numerous trials using cytotoxic regimens, particularly cyclophosphamide, and the regimen pioneered at the NIH has become widely accepted for certain types of severe renal lupus. Other internal organs can be involved with the disease, such as the lungs, heart, and CNS, and these can often dominate the clinical picture. Treatment for mild to moderate disease still relies heavily in the short term on symptom suppression by oral corticosteroids, but eventual widespread and substantial morbidity may well be due as much to the deleterious effects of long term steroids as to the disease itself.

The epidemiology and natural history of lupus have long suggested a hormonal component, and this was part of the motivation for exploring whether androgenizing hormones might be beneficial in lupus. DHEA is a naturally occurring steroid produced principally by the adrenal glands, and its metabolite, DHEA-S is the most abundant circulating adrenal steroid in humans. Women with lupus have been observed to have reduced levels of DHEA and DHEA-S, and the NZB/NZW murine lupus model demonstrates improvement with androgen treatment.

2.1 Other trials of DHEA in SLE – described in Appendix, end of review:

2.1.1 RCT (van Vollenhoven #1): Comparison of 200mg DHEA and placebo in mild to moderate SLE (Arthritis Rheum 38: 1826-31, 1995)

2.1.2 RCT (van Vollenhoven #2): Comparison of 200mg DHEA and placebo in severe SLE (Lupus 8:181-187, 1999)

2.1.3 Taiwan RCT (GBL96-01): Comparison of 200mg DHEA and placebo in mild to moderate SLE.

2.1.4 Open label study (van Vollenhoven #3): Treatment of SLE with DHEA: 50 Patients treated up to 12 months (J Rheum 25: 285-9, 1998)

2.1.5 Open label PK/clinical study (van Vollenhoven #4): Escalating doses from 50mg/day to 600mg/day of DHEA in females with SLE (J Rheum 25:2352-6, 1998)

2.2 Human Pharmacology, Pharmacokinetics, Pharmacodynamics – see pharmacology review

There is one pharmacology study of particular clinical interest, Study GL96-02, designed to evaluate the effect of DHEA on the PK of orally administered prednisone (see Pharmacology Review). This study investigated both the endogenous pituitary-adrenal axis (by ACTH stimulation test and 24 hr urinary cortisol) and the PK profile after a single 20mg dose of prednisone in 14 premenopausal normal female volunteers timed to the follicular phase of the menstrual cycle, before and immediately after a 28 day exposure to DHEA 200mg/d. Mean prednisone and total and free prednisolone were equivalent for Cmax and AUC at these two time points by the usual PK bioequivalence standard (95% confidence intervals were all within the 80%-125% window). Mean 24 hour cortisol levels were described as "unchanged from baseline" (data not included).

There was noticeable blunting in some patients of the ACTH stimulation response after the 28-day exposure to DHEA compared to pre-exposure ACTH responses. It is unclear whether this can be reasonably attributed to the design (including measurement variability/error), or whether this is a physiologic finding in certain patients. If the latter is true, it becomes necessary to access whether this is clinically significant. An external Endocrine Consult was requested to further evaluate these findings.

3. Clinical Studies

3.1 Indication #1

3.1.1 Trial #1 - Study GL94-01: Comparison of 100mg/day and 200mg/day of DHEA with placebo (PLC) in women with mild to moderate lupus.

Enrollment Dates: First patient enrolled=June 6, 1994
Last patient finished= May 28, 1996

3.1.1.1 Objective/Rationale/Primary Hypothesis

Study GL94-01 was an efficacy and dose-comparison trial designed primarily to assess durable steroid sparing of DHEA. This goal was defined in two ways:

(1) To maintain uninterrupted DHEA therapy for at least seven months, with steroid reduction sufficient to attain physiologic levels (defined as 7.5mg/d or less of prednisone or its equivalent) for at least two months at trial end. Criteria for the reduction in steroid doses for patients stabilized or improved were specified in the protocol. The proportion of patients achieving this goal in the two groups would be compared.
(2) To reduce prednisone generally (to be evaluated by comparing mean % reduction in prednisone dose in the treatment and control groups) .

Reviewer's Comment: Conception of the design

Study GL94-01 was intended to be the first of two pivotal trials to demonstrate the efficacy of DHEA in mild to moderate lupus. A small, but randomized and controlled, pilot study had been done at Stanford University (see Review Appendix – van Vollenhoven trial #1) which gave a clear signal of drug activity at 200mg/day. A formal efficacy trial was the logical next step, including a dose comparison because the Stanford study used only 200mg/d.

For RCT design, the rapid onset efficacy of steroids in mild to moderate lupus means one must either hold steroids constant, which is difficult clinically, or use steroid sparing itself as the primary outcome. The latter was elected in this trial; the endpoints were explicitly designed to directly reflect steroid increases or decreases. The design, consequently, mitigates against showing an effect in disease activity (signs and symptoms), because improvement would be offset by steroid reduction dictated by the protocol. The two primary endpoints were thus: (1) attaining physiological prednisone levels for two months at month seven or beyond and (2) reducing the mean prednisone dose.

The dual primary endpoint approach was used because of the sponsor's desire to secure Subpart E status, (an agency provision to enhance agency interaction with sponsors, in order to accelerate the development process of treatments for severe or debilitating conditions) which required a clinically important

endpoint. Thus, the first endpoint was formulated as described, and the second, as noted above, was to measure the mean reduction in steroid dose needed. This latter endpoint seemed the most straightforward for a steroid sparing design, and should be expected to be more discriminating to detect an effect, if present. Although power calculations were done for both endpoints, powering decisions were made on the basis of the second endpoint, including aspects related to the planned interim analysis (see Section 3.1.1.3.3, below).

3.1.1.2 Design

This was a parallel arm, 7-9 month, randomized, double-blind, trial comparing 100mg/day, 200mg/d of DHEA, and placebo in patients with mild to moderate lupus whose disease had been resistant to steroid reduction. The primary analyses were (1) a comparison of patients able to reduce steroids to physiological levels for at least two months including months 7, 8, or 9, and (2) a comparison of the mean steroid reduction across arms.

Reviewer's comment: The trial duration was variable – 7 to 9 months -- because of the desire to exit patients from the trial at month seven if they had already satisfied the endpoint requirement for treatment over two months at physiological doses of steroids. There was no other clinical need for this design feature. Patients were given the option of enrolling in the open extension, Trial 95-01. No other formal follow-up was provided in the design.

3.1.1.3 Protocol

3.1.1.3.1 Population, procedures

For entry, women needed to be a) at least 18 years old with a diagnosis of SLE by ACR criteria (see appendix of protocol), b) needed mild to moderate disease characterized by a prednisone requirement of at least 10mg/day but not more than 30mg/day, and c)they needed to be seriously steroid dependent ("steroid stuck") as evidenced by either (i) a failed attempt to taper prednisone within the last 12 months, and stable in the last 6 weeks, or (ii) no prednisone taper failure but a stable dose in the last 3 months. Nonsteroidal anti-inflammatory drugs (NSAIDs) and/or hydroxychloroquine doses must have been stable for the prior one month. Use of cytotoxic medications was an exclusion criterion. Patients of child-bearing potential must have had a negative pregnancy test within two weeks of entry and be using a reliable, non estrogen-containing form of birth control.

The trial used an algorithm to prescribe steroid usage as a function of the clinical activity as measured by the SLEDAI score. This was a critical feature of the design, as the goal was to capture the ability to reduce steroids where they were not needed – i.e. if the patient was stable. In particular, the following were specified in the protocol:

For all patients with an unchanged or improved SLEDAI:

If the daily prednisone was:	The dose needed to be reduced by:
>0 - ≤5mg	1mg/day
>5mg - ≤10mg	2.5mg/day
>10mg - ≤30mg	5.0mg/day
>30mg	to be determined by physician

Reviewer's Comment: Steroid-dependency in lupus & its implication for assessment:

In the design deliberations, there was extensive discussion on the best way to capture serious steroid dependent ("steroid stuck") patients. That such patients clearly exist was not in question, but this issue complicates the design and conduct of many lupus trials. (1) Some lupus patients tend to self-medicate, changing their steroids on their own. (2) There is a wide variability among practitioners in the manner in which steroids are used in lupus, from very liberal to very conservative. (3) There is a tendency in some lupus patients to develop a dependency on steroids, taken as much to avoid steroid withdrawal symptoms (adrenal insufficiency) as for treating lupus per se. (4) In many patients with mild lupus there is an intermingling of constitutional symptoms (e.g. fatigue, some thinking this more due to a concomitant, but etiologically distinct, fibromyalgia), psychological symptoms (subtle CNS lupus or reactive, but etiologically distinct, symptoms), and steroid-related symptoms (induced, or secondary to withdrawal).

3.1.1.3.2 Endpoints

This study employed two primary endpoints. The first, construed for the purposes of pursuing Subpart E status, consisted of completing at least seven months of treatment for at least two months at trial end at “physiological doses” of prednisone (7.5mg/day or less). These patients were called "responders". The second primary endpoint was a simple comparison of mean prednisone reduction.

Reviewer’s Comment: The rationale for these primary endpoints is given above (Sections 3.1.1.1)

3.1.1.3.3 Statistical considerations

PRIMARY ANALYSIS

The first primary analysis was a comparison of responders by logistic regression, with the model including any baseline covariate attaining 0.05 significance level for association with treatment assignment. A modified analysis was also proposed (Amendment 5, dated March 21, 1997) adding baseline SLEDAI and treatment interaction to the model.

The second primary analysis was a comparison of the percent decrease in prednisone from baseline by one-way ANOVA model, with use of any baseline covariate attaining 0.05 significance level for association with treatment assignment, and a Bonferroni adjustment for multiple comparison (100mg/day vs. placebo and 200mg/day vs. placebo). The same modified analysis was similarly proposed here (Amendment 5, dated March 21, 1997), adding baseline SLEDAI and treatment interaction to the model.

SECONDARY ANALYSIS

The secondary analyses were comparisons of mean changes from baseline in the following: SLEDAI, Krupp Fatigue Severity Score (KFSS), patient global (PG), and investigator global (IG), by a one-way ANOVA with baseline value as a covariate, including treatment-by-baseline interaction included in the model, and a Bonferroni adjustment for multiplicity.

> ***Reviewer's comment: As noted above, although the design did not exclude the possibility of showing treatment effects in these parameters, the primary goal of decreasing steroid dose diminished the likelihood of demonstrating a treatment effect upon the secondary endpoints. Therefore, these secondary analyses were not accorded, a priori, any inferential weight regarding drug attribution. In other words these analyses were not expected to succeed.***

ANALYSES PROPOSED IN AMENDMENTS (AS NOTED ABOVE)

A comparison of responders by logistic regression, using baseline SLEDAI as a covariate (Amendment 5, dated March 21, 1997).

A comparison of responders by logistic regression of the SLEDAI>2 subset.

> ***Reviewer's comment: Before unblinding it became evident that the responder rates were higher in patients with low baseline SLEDAI scores (SLEDAI 0-2: 65%, SLEDAI 3-4: 43%, SLEDAI 5-8: 43%, SLEDAI>8: 31%), and that this might dilute a drug effect if this high rate was equally distributed across all three groups. A decision was made to conduct a SLEDAI>2 subset analysis, and to increase sample size of the ongoing second trial (Trial GL95-02, see below) with SLEDAI>2 patients, in the hopes that this would increase power by increasing numbers and enriching with SLEDAI>2 patients.***

SAMPLE SIZE CONSIDERATIONS

There was no prior information to formally power this trial. At an alpha of 0.05, 190 patients, reduced to 168 (56 per arm) after withdrawals, yielded approximately an 80% power for the pair-wise, two-sided tests of the first primary efficacy (achieving physiological steroid doses) if one arm showed a 5% and the other a 22% response rate. The power is also about 80% for outcomes of 10%/30% or 20%/44%. The power is about 50% if the rates are 5%/16%, 10%/24%, or 20%/37%. With respect to the second

primary efficacy (mean change in prednisone use), 56 patients per arm yields an 80% power for the pair-wise, two-sided tests, if the difference between treatments is assumed to be 30% and the within-group, between-patient standard deviation is 55. The assumption of a standard deviation of 55 was provisional, and could be adjusted by use of the blinded interim analysis estimating the pooled variability, in order to ensure 80% power for a treatment effect of 30%. The interim analysis would be done after 60 evaluable (i.e. having a baseline and at least one post-baseline visit data) patients were available.

> ***Reviewer's Comment: After some discussion of what exactly what would be the inferential implications of the planned interim analysis (eventually leading to amendments 4 and 5, attached), the sponsor decided that an interim analysis would need a multiplicity adjustment of approximately 0.5, i.e. a two-sided, pair-wise comparison at the level of 0.025.***

HANDLING OF WITHDRAWALS

Withdrawals were not explicitly addressed in the protocol (except in the power calculations), but success by the first primary endpoint required at least seven continuous months on therapy. The protocol only stated that the analysis of the first primary endpoint would be by logistic regression of the proportion achieving success, and that of the second primary endpoint would be by ANOVA.

> ***Reviewer's Comment: Missing data in any RCT can bias the result, leading to a false positive outcome if a withdrawal pattern favored the drug or a false negative outcome if it favored the control, but here both primary efficacy variables failed to show statistical significance, so the issue of bias favoring DHEA need not be formally explored.***

> ***Reviewer's Comment: Given the generic problem in prior rheumatology trials related to interpretation of withdrawals, the endpoint definition in this trial was carefully designed to mitigate this. The simplest way to do this was to call a "success" not only one who met disease criteria, but who also met durability criteria --here having completed at least seven months on therapy. Thus, a patient could not have been a withdrawal – for any cause – before month seven and still be counted as a "success" (i.e. in the numerator) in the first primary analysis. One could perform a simple Fisher's exact comparison of proportions here. However, to enable covariates to be incorporated in the analysis of the first endpoint, one needed a logistic regression analysis, which as conventionally done is time-anchored and thus needs imputation which traditionally has been the LOCF method. The second endpoint was assessed using an ANOVA that allows for covariate adjusting. It too used LOCF for imputation.***

3.1.1.4 Results

3.1.1.4.1 Patient accountability

	DHEA200	DHEA100	PLC	TOTAL
Total randomized	64	63	64	191
Completed 7-9 months	47	46	49	144
Withdrawals	17	17	15	49
Reason:				
Inefficacy	5	6	7	18
Adverse event	6	4	3	13
Other	6	7	5	18

Analysis by survival methods for withdrawals – all reasons, inadequate efficacy, and adverse event – did not show any statistically significant differences across any arms, although this analysis obviously had limited power. The survival curves are displayed in Table 2 (3 pages, appendix).

According to investigator attribution, 6 patients were withdrawn for an event possibly related to DHEA, five on 200mg/day (headache/nausea/backpain, decreased WBC/elevated LFTs, sores on buttocks, worsening rash/alopecia/pruritis/worsening purpura, facial dermatitis), and one on 100mg/day (hirsutism/acne).

3.1.1.4.2 Patient comparability at entry

DEMOGRAPHICS	DHEA200	DHEA100	PLC
Age (mean, range)	40 (21-66)	40 (18-75)	41 (22-70)
Race - %w/b/other	55/27/18	57/25/18	69/27/4
Pre-menopause	75%	59%	59%
Pre-exist hypertension	55%	37%	36%
Current smoking	45%	56%	45%

BASELINE STATUS	DHEA200	DHEA100	PLC
Prednisone (mg)	13.7	13.7	15.2
SLEDAI (range)	5.9 (0-22)	5.3 (0-16)	6.4 (0-22)
KFSS	5.4	5.1	5.3
PG (10cm VAS)	47	46	49
IG (10cm VAS)	23	26	28
SF-36 physical	32	35	33
SF-36 mental	45	45	43
SLICC damage index	2.3 (0-9)	2.5 (0-13)	2.1 (0-9)

None of the demographic or baseline comparisons with data above demonstrated statistically significant differences (by Cochran Mantel-Hansel Chi-Square test.)

Reviewer's Comment: The SLEDAI measure played an important role in both this and the subsequent trial (GL95-02). The distribution of the baseline SLEDAI values is of interest. These data are shown below in the bar graphs of Table 3 (appendix), and summarized numerically.

	--------number of patients---------		
	DHEA200	**DHEA100**	**PLC**
SLEDAI			
0-2	**19**	**16**	**19**
3-4	**9**	**18**	**15**
5-8	**24**	**16**	**13**
>8	**12**	**13**	**17**
total	**64**	**63**	**64**

3.1.1.4.3 Efficacy endpoint results

3.1.1.4.3.1 Primary analyses per original protocol

FIRST PRIMARY ANALYSIS: Comparison of responders -- completing trial and achieving physiological dose steroids for at least two months at end of trial -- by logistic regression. As specified in the protocol, all baseline variables (see **Section 3.1.1.4.2)**, including SLEDAI domains and specific renal components (hematuria, proteinuria, pyuria) and all SF-36 domains were analysed for baseline imbalance, and none were found to be statistically significant. Thus, the primary analysis, P values for pair-wise comparisons of the DHEA groups with placebo, was done unadjusted.

Result:	DHEA200mg	DHEA100mg	PLC
	35/64	28/63	26/64
p value	0.11	0.66	

SECOND PRIMARY ANALYSIS: Comparison of mean percent change in prescribed prednisone dose by one-way ANOVA model, again with no covariate adjustment. P values are for pair-wise comparisons of the DHEA groups with placebo.

Result:	DHEA200mg	DHEA100mg	PLC
	-30.3% (+/-74.3)	-13.7% (+/-91.4)	-35.8% (+/-50.2)
p value	0.67	0.09	

To further display the dynamic occurring here with baseline and changing prednisone during the trial, scatterplots are provided in the appendix (Table 4, 3 sheets)

Reviewer's Comment: As noted earlier, one might expect the comparison of mean change in prednisone to be the more sensitive (i.e., discriminating) of the two primary endpoints. It measures directly the factor relevant to the

hypothesis, steroid sparing. The responder test, in contrast, is a dichotomous instrument, and these are known to always engender loss of information. Additionally, it is derivative in the sense of being a number of steps removed from the phenomena impacted by therapy – signs and symptoms. One needs to first assume that the SLEDAI does, in fact, well capture treatment changes in signs and symptoms, then that the steroid dose titration is executed as intended, and finally that the patient continues to perform adequately and achieves steroid reduction to 7.5mg/day of prednisone in order to "win" by the first endpoint.

3.1.1.4.3.2 Secondary analyses

As noted above, these secondary analyses are meant for descriptive use only, because these variables were NOT EXPECTED TO CHANGE in the trial due to the design (see **Section 3.1.1.3.2***, above):*

P-VALUES FOR GROUP MEAN COMPARISONS TO PLACEBO

	DHEA200	DHEA100
SLEDAI	0.75	0.38
Pt. Global	0.37	0.28
Inv. Global	0.66	0.93
Krupp Fatigue scale	0.96	0.77
SF-36 Physical	0.99	0.62
SF-36 Mental	0.35	0.76

3.1.1.4.3.3 Primary Analyses incorporating modifications to original analysis plan

Comparison of responders by logistic regression using baseline SLEDAI as a covariate (Amendment 5, March 21, 1997).

Result:	DHEA200mg	DHEA100mg	PLC
	35/64	28/63	26/64
p value	0.12	0.84	

Reviewer Comment: As noted above, by the end of trial GL94-01 the sponsor drew the conclusion that baseline SLEDAI was a predictor of DHEA response. It is therefore understandable that the sponsor would amend the protocol in this way. Although the protocol would have allowed an adjustment for baseline SLEDAI if it were found to be significantly imbalanced ($p<0.05$), this amendment would allow the adjustment, whether or not there was a statistical imbalance at baseline. As it turned out, the eventual analyses, adjusted for baseline SLEDAI, differently insignificantly – the DHEA200 v PLC p value increased from 0.11 to 0.12, and the DHEA100 v PLC p value increased from 0.66 to 0.84.

3.1.1.4.3.4 Exploratory analyses of responsiveness of subsets

A. RESPONSE BY SLEDAI >2 SUBSET: The sponsor also reported analyses for the subset defined as entry SLEDAI >2. This subset consisted of 137 patients. A clinical argument was also given, asserting that patients of low SLEDAI scores frequently had their scores based on laboratory abnormalities such as complement or DNA levels, which often do not change even when symptoms do. In the case of Trial 94-01, 54 patients had a baseline SLEDAI score of two or less, including 26 with no disease activity (SLEDAI=0) and 20 with SLEDAI of 1 or 2, due only to serologic abnormalities (elevated dsDNA antibody or decreased complement). No minimum SLEDAI score was required for entry.

The number of patients with various components of the SLEDAI for those entering this trial with a SLEDAI score of 2 or less are shown below:

	DHEA200	DHEA100	PLC
Increased dsDNA	7	5	6
Low complement	0	1	1
Mucosal ulcers	1	1	0
Alopecia	0	1	0
New rash	0	1	0
Leukopenia	1	0	0
Pleurisy	1	0	0

The results of the logistic regression analysis of the SLEDAI>2 subset, using the same covariate criteria as above, are shown.

Result:	DHEA200mg	DHEA100mg	PLC
	23/45	18/47	13/45
p value	0.18-adj*	0.75-adj*	

* adjusting for baseline prednisone, the one covariate which proved imbalanced (p=0.039) at baseline in this subset. The sponsor also provided an unadjusted analysis, which yielded a p value of 0.031 for the comparison of DHEA 200 mg to placebo.

B. RESPONSE BY ALL SLEDAI SUBSETS:

		----------proportion responders----------		
		DHEA200mg	DHEA100mg	PLC
SLEDAI	n			
≤2	54	12/19 (63%)	10/16(62%)	13/19 (68%)
>2-4	42	5/9 (56%)	8/18 (44%)	5/15 (33%)
>4-8	53	12/24 (50%)	7/16 (44%)	4/13 (31%)
>8	42	6/12 (50%)	3/13 (23%)	4/17 (23%)

		----------mean percent (SD) prednisone change----------		
		DHEA200mg	DHEA100mg	PLC
SLEDAI	n			
≤2	54	-50% (65.3)	-54% (31.9)	-60% (29.1)
>2-4	42	-47% (27.3)	-23% (61.5)	-20% (72.7)
>4-8	53	-18% (86.2)	-10% (87.2)	-30% (41.0)
>8	42	-11% (82.9)	44% (145.5)	-27% (45.6)

C. RESPONSE BY STEROID HISTORY AT ENTRY

Patients could qualify for this trial (see **Section 3.1.1.3.1**) if either (1) they had a failed prednisone taper within 12 months, or (2) they had stable steroids for at least three months. The responses in each of these subsets is shown below.

	----------proportion responders----------		
	DHEA200mg	DHEA100mg	PLC
entry criterion			
stable steroids	5/10 (50%)	4/10 (40%)	6/16 (37%)
failed taper	30/54 (56%)	24/53 (45%)	20/48 (42%)

Reviewer's Comment: The database is too small to identify which of the two criteria for defining steroid dependence is the more valid or sensitive to treatment effect.

D. RESPONSE BY DOMINANT ORGAN INVOLVEMENT AT ENTRY

Reviewer's Comment: This area was not explored, as it would have involved dissecting apart the individual components of the SLEDAI, or by chart reviews to categorize patients by the predominant organ involvement at trial entry. It would be worthwhile exploring.

E. RESPONSE BY RACE AND MENOPAUSAL STATUS

	---------- proportion responders (%) ----------		
	DHEA200mg	DHEA100mg	PLC
Race:			
African-American	47%	44%	35%
Caucasian	54%	47%	43%
Menopausal status			
Pre-menopause	54%	38%	42%
Post-menopause	43%	47%	37%

F. RESPONSE BY BASELINE LABORATORY VALUES:

One can also evaluate whether a baseline laboratory value (e.g. DHEA-S, or estradiol), or a treatment-induced change of a laboratory value, correlate with response. The NDA included scatter plots of these analyses for baseline and on-treatment DHEA-S, estradiol, testosterone, and HDL-cholesterol, but no clear patterns emerged. The data are reproduced in Table 5 (9 sheets, appendix). It is important to note that blood drawing was not timed to DHEA administration, potentially contributing to variability which could blur any relations which might exist.

(DHEA-S levels were done in this trial, although, as noted above, these were not timed to drug administration. For descriptive purposes, the levels are shown in Table 6 below.)

TABLE 6: BLOOD LEVELS OF DHEA-S IN TRIAL GL94-01
Figures are number of patients

	DHEA200			DHEA100			PLC		
	pre-rx	**first**	**last**	**pre-rx**	**first**	**last**	**pre-rx**	**first**	**last**
DHEA-S (ug/dL)									
<250	**60**	**14**	**20**	**61**	**11**	**25**	**61**	**54**	**53**
250-500	**1**	**5**	**11**	**0**	**12**	**12**	**0**	**1**	**0**
500-1000	**0**	**11**	**11**	**0**	**24**	**16**	**0**	**0**	**0**
1000-2000	**1**	**14**	**12**	**0**	**9**	**8**	**0**	**1**	**0**
>2000	**1**	**17**	**9**	**0**	**3**	**1**	**0**	**0**	**0**

3.1.1.4.4 Safety comparisons

3.1.1.4.4.1 Extent of exposure

	DHEA200mg	DHEA100mg	PLC
N	64	63	64
Mean (day)	177	175	172
Medium (day)	194	195	196
SD (day)	46	50	55
Range (day)	12-224	7-232	2-236

3.1.1.4.4.2 Adverse events

EVENTS OCCURRING IN AT LEAST 10% OF PATIENTS: Table 7 (2 sheets, appendix) shows adverse events occurring in at least 10% of patients, sorted by the occurrence rate in the 200mg/d patients. Acne, abdominal pain, and hirsutism were more common in the DHEA cohorts.

EVENTS CONSIDERED "SEVERE" REGARDLESS OF RATE OF OCCURRENCE OR OF ATTRIBUTION: Table 8 (3 sheets, appendix) shows all adverse events assessed as severe, without regard to attribution. Most categories are small number instances.

ADVERSE EVENTS AS A FUNCTION OF SLEDAI SUBSETS: This was an attempt to see if there was a preferential DHEA effect on certain SLEDAI subsets. The sponsor was requested to tabulate all adverse events by body system, subdivided by the four baseline SLEDAI groups used above. These data are supplied in the appendix as Table 9 (14 sheets). There did not prove to be notable tolerability differences across SLEDAI subsets.

DEATHS: No deaths were reported during the study, but three were reported in the follow-up period. One PLC patient had an abdominal abscess and severe anemia, but she refused transfusion and died of respiratory failure and massive bleeding. A second on DHEA100mg had respiratory failure secondary to lupus and a thrombotic microangiopathy, and the third on DHEA200mg had hemorrhagic pancreatitis. The time of death in these patients was 2, 3 ½, and 9 months, respectively, after discontinuing study drug.

MALIGNANCIES: None reported during this trial, but see **Discussion (end of review).**

ABDOMINAL PAIN: There were a total of 18 patients on DHEA200mg with abdominal pain or related symptoms, 10 on DHEA100mg, and 6 on PLC. A line listing of these patients is in the appendix as Table 10 (2 sheets). Although this "signal" regarding GI events associated with DHEA was noted, analysis of the by-patient information reporting nausea and vomiting and/or abdominal pain defied any obvious common pathophysiology.

HYPERTENSION: The following is a table of patients with new onset or worsening hypertension during this trial.

	DHEA200mg	DHEA100mg	PLC
new onset HT	11	11	11
increased HT	19	4	18
increased rx	5	7	0

Of those patients with a prior history of, or pre-existent, hypertension, 11 of 23 on DHEA200mg developed hypertension, 5/9 on DHEA100mg, and 7/9 on PLC.

Blood pressure was further analyzed, as there was concern regarding possible increases in blood pressure on DHEA because of its being metabolized to testosterone. The mean and median DBP and SBP did not change throughout the trial in any arm:

	Mean (median) values	
DHEA200mg	DHEA100mg	PLC

DBP-Initial	80.1 (80)	78.5 (78)	79.6 (80)
Final	79.2 (78)	75.8 (78)	76.6 (80)
SBP-Initial	126.7 (124)	124.7 (120)	121.9 (120)
Final	122.4 (120)	118.2 (120)	121.1 (120)

WEIGHT: There was no difference in weight over the trial in any arm.

3.1.1.4.4.3 Laboratory Data

3.1.1.4.4.3.1 Descriptive statistics

HORMONE LEVELS:

Dose related changes were observed in serum estradiol and testosterone (increased) and in gonadotropins (decreased), consistent with the suspected endocrinologic impact of DHEA. These data are referred to below for the premenopausal (n=121) and physiologically-defined, post-menopausal (n= 40) subsets. (The remaining 28 patients had various forms of surgical menopause, and their data are supplied in the NDA). The visits were not timed to the menstrual cycle.

ESTRADIOL: There was a considerable increase in estradiol associated with DHEA treatment in those post-menopausal patients not on exogenous hormone replacement therapy (see Table 11, below), but the number of patients was only four. Nonetheless, because this was consistent with the anticipated phamacology of DHEA, and because of concern regarding possible deleterious effects of prolonged, unopposed estrogen on breast and endometrial tissue (See Endocrine Consultation Review), a program to monitor these post-menopausal women (not on hormone replacement therapy) in the ongoing trial #2 (GL95-02) with uterine ultrasound and biopsy as needed plus mammography was instituted. To date, there are 43 post-menopausal patients who have been studied with mammography under this program, and 24 with uterine ultrasound (see **Section 3.1.2.4.4.2**).

> ***Reviewer's Comment: The question of long term sequalae to hormonal DHEA exposure has been present throughout the development, and it will impact the configuration of the overall safety database.***

FSH/LH: Only the post-menopausal group showed a significant decline, as noted in the Table 12 (appendix).

TESTOSTERONE: Testosterone levels also show a dose-response, more pronounced in the post-menopausal patients; the data are shown in Table 13 (appendix).

LIPID LEVELS: The values for total, HDL-, and LDL-cholesterol levels and the ratio of cholesterol/HDL-cholesterol are shown in Table 14 (5 sheets, appendix). Sporatic between-group comparisons reached statistical significance, and some may be an effect secondary to an increase in testosterone . Interpretations of these data are complicated by the small size of physiologically defined subgroups, the known confounding by prednisone doses which were not kept constant, and the known abnormal lipid profiles de novo in many lupus patients. However the differences seen in table 14 are of potential concern and warrant further study. Any increased risk in atherosclerotic disease would need to be included in a risk benefit assessment.

OTHER LABORATORY VALUES: A survey of all laboratory values, comparing mean changes over the trial duration, is given in Table 15 (41 pages, appendix).

COMPLEMENT: The values for complement levels, C3 and C4, are shown in Table 16 (appendix). A trend of ***lowering*** of both was observed, reaching statistical significance in the 200mg DHEA group for C4. This signal is of concern and warrants further study.

URINARY PROTEIN: This parameter showed a dose-dependent increase in proteinuria in the DHEA groups. The mean changes in 24-hr urinary protein were: - 34mg/d in PLC, compared to +323mg in DHEA100mg and +870mg in DHEA200mg. Statistical tests of these changes were not provided in the NDA. Baseline and last visit urinary protein values are shown in scatterplots form in table 17 (6 sheets, appendix).

> ***Reviewer's Comment: It was the trends noted in complement levels, and the urinary protein changes that lead to the following (prospective) exploratory analysis by the reviewer (see Section 3.1.1.4.4.3.2).***

3.1.1.4.4.3.2 Reviewer analysis of patients displaying worsening in parameters associated with renal lupus

The above signals suggesting falling complement and rising urinary protein lead to the question of whether DHEA was inclined to exacerbate (or initiate) renal lupus. Using pre-specified criteria and well-accepted laboratory conventions, the question of whether there was a worsening of parameters of urinary sediment or serology was explored, given that these could be interpreted as markers for possible new or worsening renal involvement by lupus. These criteria are listed in the table below. Urinary protein, hematuria, complement levels (C3 and C4), and double-stranded DNA antibody titers were the parameters used. The definition of worsening for any parameter was set to require a second (i.e., follow-up) laboratory result similarly abnormal as confirmatory. This criterion which would cause some true cases to be missed ("false negatives"), but it would enhance the veracity of cases so selected ("true positives"). The only exception to needing a

confirmatory value was if, after one abnormal value, the patient was dropped out of the trial specifically for a worsening of that parameter. However in neither GL94-01, or in GL95-02 were there any such patients. Then, the full laboratory trial experience for all patients was reviewed to determine the number of patients who proved to have evidence, as above defined, for worsening in these four parameters.

DEFINITIONS OF ABNORMALITIES SUGGESTIVE OF RENAL LUPUS

Parameter	if normal at baseline	if abnormal baseline
Urinary RBC count (nml: 0 RBC/hpf)	increase to $\geq$10	double to $\geq$10
24 hr urinary protein (nml $\leq$150mg/d)	double to >150mg/d	double
complement C3 (nml: 85-193)	25% decrease	25% decrease
complement C4 (nml: 12-36)	25% decrease	25% decrease
anti-double-stranded DNA (nml: 0-3.6)	double	double

RESULTS:

Patient number

	urine RBC	urine protein	falling C3	falling C4	rising DNA
PLC	12120	20222	15164	12115	12115
	16194			15294	14112
				18216	14228
				20155	17175
					20155
					23257
DHEA	3153	14107	17332	19132	12113
100mg	15161	17332	18217		12116
	18284	18128			13102
	20225	18137			14108
	21200	18282			17332
					18217
					23190
DHEA	14111	18140	3149	3149	13106
200mg	15234	18143	13103	18140	14227
	18143	18318	14111	18142	18219
	18145	20156	14227	18219	
	20156		15236	19134	
	22172		17173	21197	
			18140		
			21335		
			24275		

If one determines the number of patients (no patient counted twice) with at least one laboratory test suggesting new or worsening lupus disease, the following are obtained.

PLC	n=12/17	(19%)
DHEA 100mg	n=17/63	(27%)
DHEA 200mg	n=20/64	(31%)

The number of patients with more than one such test positive are:

PLC	n=2
DHEA 100mg	n=2
DHEA 200mg	n=7

Reviewer comment: These results suggest a "signal", and merit assessing in GL95-02 (see Section 3.1.2.4.4.3.2), If similar findings are seen there, then this will need to be further explored.

3.1.1.5 Conclusions:

1. **The drug (200mg, 100mg, respectively) versus placebo comparisons for the first primary endpoint were p=0.11 and p=0.66, respectively, and for the second primary endpoint, 0.67 and 0.09, respectively.**
2. **Exploratory analyses suggest a possible effect on patients with SLEDAI >2.**
3. **Adverse events and relevant lab changes occurring with a higher rate in DHEA treated subjects included abdominal pain, hirsuitism, acne, HDL levels, proteinuria, hematuria, change in C3 and C4.**
4. **Levels of estrogenic and androgenic hormones followed patterns anticipated with the administration of a hormone precursor with the metabolic pathways known for DHEA**

3.1.2 Trial #2 - Study GL95-02: Comparison of 200mg DHEA and placebo in mild to moderate SLE

Enrollment: First patient enrolled=March 7, 1996
Last patient finished=April 2, 1999

3.1.2.1 Objective/Rationale/Primary Hypothesis:

RELATION OF DESIGN TO OVERALL NDA RATIONALE: This trial was designed to assess the signs and symptoms of SLE in patients treated with 200mg daily of DHEA over a one-year period. The goal was to provide a demonstration of improvement in disease activity, which, then, along with Trial GL94-01 demonstrating steroid sparing, could serve as adequate evidence of efficacy of DHEA in the treatment of lupus. A closed Arthritis Advisory Committee (March, 1995) was convened to review this strategy and, in principle, concurred, provided other elements were satisfactory. Chief among these other elements were an assurance that DHEA did not simply enhance the PK or PD effects of prednisone, some data regarding its use in severe lupus and in men, and an adequate safety profile for risk / benefit. Since steroid sparing was part of the claim rationale, rigorous demonstration of the absence of any such prednisone enhancement would offer some, but not complete, assurance that the long-term safety database for DHEA would not prove to replicate or mimic that which exists for the long-term steroid use. True risk / benefit of any steroid sparing effect would be contingent upon adequate long-term safety data on DHEA.

Consequently, there was extensive communication between the sponsor and the agency regarding the design of Trial GL95-02, how it linked with the design of Trial GL94-01, and how both might constitute an NDA strategy. The sponsor summary of these considerations is attached to this review as Table 18 (7 sheets, appendix).

Reviewer's comments: Conception of the primary hypothesis

Key to any RCT design in lupus is use of credible instruments to assess efficacy, but this fundamental precondition is itself problematic in lupus. There is no simple, established disease assessment available. Also, there are virtually no design precedents in lupus to use to judge the performance of candidate instruments. Because of these clinical realities, a careful articulation of the primary hypothesis, a clear agreement that the hypothesis is relevant to patients and so usable in drug approval, and a prospective, detailed description of how the hypothesis being tested translates into the statistical analysis, were all deemed critical. The sponsor, agency, and both groups of consultants were aware of the paucity of precedent here and the inherent uncertainty about the adequacy of any instrument. It should therefore be noted that a failed study could result from either drug failure or design failure.

3.1.2.2 Design: Study 95-02 used a double-blind, placebo-control, two-arm, one-year, randomized design. The primary endpoint was a by-patient, dichotomous measure of response. The primary analysis is a comparison of the proportion of responders in the two arms, with prespecified covariates (race, cytotoxic use, prednisone use, baseline values of the four components of the responder definition (see below), and menopausal status) assessed for inclusion. A sub-study to investigate the effect of DHEA on bone density was specified.

Reviewer's comments: Deliberations Relevant to Determination of the Primary Endpoint

#1-General comments:

As with all RCT designs, known and suspected risk factors and confounders needed to be addressed in the design or in the analysis. This means, typically, a careful selection of entry criteria to avoid stratification or enrollment imbalances for known risk factors. There was forethought here. For example, patients needing daily prednisone of more than 10mg/d or any changing hydroxychloroquine or cytotoxic therapy in the prior six weeks were excluded. The greater design challenge here, as in Trial GL94-01, with patients with mild to moderate lupus – who can often get dramatic, short-term symptom relief with an increase in their daily steroids, was how to prohibit these obviously confounding interventions. In fact, it was in deference to this fundamental of lupus management that the sponsor's first trial (GL94-01) accepted the clinical reality of steroid dependency, and incorporated it as its primary hypothesis of steroid sparing.

In trial GL95-02 strict provisions on steroid use were employed. Specifically, during the first two months (the sponsor's estimate of DHEA time-to-onset) up to 10mg/d increases were allowed. For the remainder of the trial up to 5mg/d increases for two months maximum were allowed.

#2-Primary endpoint formulation and rationale

As mentioned above, there is no heritage or consensus on measurement of the status of the lupus patient, nor even what concepts -- signs/symptoms/quality of life/psychological state/irreversible damage -- should be captured in such a determination. In light of this complexity and attendant unknowns, it was decided at the design stage that a strategy which provided for the determination of a responder or non-responder status would be the least controversial approach.

Clearly, some concept of disease activity and the perception of the patient ("patient global") needed to be included in this composite. This was particularly true because the constitutional dimensions were thought to be under-represented in the SLEDAI and SLAM instruments used (see below). The patient's perception is particularly germane in lupus, as there is often a discordance between how the patient feels (e.g. "lousy," "tired all the time") and the physical manifestations such as rash or synovitis, which may objectively appear minor or even nonexistent.

#3-Endpoint elements to capture relevant domains in lupus

There are two competing indices of disease activity, the Systemic Lupus Erythematosis Disease Activity Index (SLEDAI) and the Systemic Lupus Activity Measure (SLAM), neither of which has been used in any RCT. In the absence of any good comparative data to choose between them, both were incorporated into the composite endpoint for this trial. Similarly, both the Krupp Fatigue Severity Scale (KFSS) and the patient global (PG) by 10 cm visual analog scale (VAS) were both used to assess the sometimes dominant constitutional aspect of the lupus patient, because there was no data to use to decide between the two. The components of the SLEDAI, SLAM, and KFSS are in the protocol appendices (attached).

As lupus is a chronic, sometimes progressive disease, certain fixed pathologic features appear in some patients over time. These aspects are referred to as "damage", and they have also been codified in an index (the SLICC index, see appendix for its components). Examples of damage are long-standing renal insufficiency or residual deficits from a cerebrovascular accident, cataracts, or old myocardial scarring from an earlier infarct. Sometimes histologic evidence of permanence is weak. In some instances the distinction between activity – presumed reversible, and damage – presumed fixed, is blurred. A relatively

recent, modest change in creatinine is an example of this uncertainty. Furthermore, damage in lupus often has its genesis in a combination of disease and drug toxicity, especially prolonged steroid use. Although damage data were collected (the SLICC was collected at baseline at endpoint or dropout) to be incorporated into the safety profile and ultimate risk benefit, DHEA was not expected to impact these features. Nonetheless, one really did not know the extent or limits of the DHEA effect, so all known aspects of lupus and serious drug toxicity were incorporated into the responder definition (see below).

#4-Endpoint features to capture other important events

Because of the possibility that the activity indices (SLEDAI and SLAM) might miss a serious clinical event which would clearly signify treatment failure, the endpoint was designed to capture all serious, irreversible events which would be considered obvious evidence of worsening disease and thus drug failure, and so included a generic criterion of instances for needing new or increased cytotoxic therapy. Thus, a responder was defined as one free of any major "clinical deterioration," defined in detail (see Section 3.1.2.3.2) during the trial or for six weeks thereafter (i.e., post-drug exposure).

#5-Endpoint feature to capture serious steroid toxicity

Constructing a "toxicity net" to capture safety events – as best they can be predicted -- is always a problem when incorporating safety into a composite endpoint. In this case this was done because of the desire for a clinically all-inclusive, by-patient endpoint, and the very real concern at the start of GL95-02 (no definitive PK work having been done to argue otherwise – see Section 2.2) that DHEA might simply have its effect via increasing prednisone exposure by making it more bioavailability, or have its effect via increasing prednisone effects by making it more pharmacodynamically available. Thus it was felt important that the endpoint "capture" severe steroid toxicity events, such as new onset diabetes, hypertension, osteoporotic fracture, and myocardial infarct.

#6-Construction of the primary endpoint:

The question of what metric would be optimal to construct the primary responder composite was thoroughly discussed, as, here, too, no precedents existed. This discussion included exactly what "cutoff" should be used for the four components. The definition of "no worsening/success"(e.g. anything better than a 10% worsening would be considered a success) would be as arbitrary as any other (e.g. 5%, or 15%). The theoretically ideal definition depends on the anticipated rates of success by the various cutoffs as well clinical relevance. The goal of the sponsor and the Agency was to define a clinically meaningful and potentially treatment-responsive endpoint definition. There needed to be enough "action" regarding the endpoint. If the chosen endpoint proved to be too rare an occurrence, it would make the design unlikely

to detect a treatment effect, even if it existed. And, if it were too prevalent, it may not be sensitive enough, and the ability of the design to be discriminating would be undermined, again leading to the risk of false negatives. Data to help make this decision did not exist. Because of this dilemma those working on the trial design decided, in the end, to simply define "no worsening" as the endpoint. Thus, improvement or no deterioration counted as a success, and ANY worsening meant a failure, as noted in the protocol language. This question was settled at the time of design, and this feature was clearly stated in the protocol used at the outset of the study. There was not any deferral of this decision into the future. What was left open as a possible problem was the caveat was that IF the endpoint did prove to be positive in, say, only 5% of one arm and 3% in the other – i.e. a very low "hit rate" overall, or if it proved to be much too prevalent, say, 90% in one arm and 95% in the other, then would there be a legitimate argument to default to alternative analyses with modifications of the endpoint definition.

Thus, a dichotomous approach was adopted, using a zero-change cutpoint for all four of the components (SLEDAI, SLAM, KFSS, PG), and a "responder" was defined as a patient who (1) does not deteriorate in any of the four components (i.e. the mean on-trial score compared to baseline for each index being zero or greater), AND (2) does not evidence any clinical deterioration criteria (detailed in Section 3.1.2.3.2).

#7-Endpoint analysis – population and statistical test

Implicit in the construction of the primary endpoint for any RCT is the analysis to be done and the statistical test (with the assumptions it carries) to be used. The protocol at the outset of the trial specified the logistic regression. The null hypothesis model mandates, to preserve the validity of causal inferences, that the (full) randomized cohort be the population for analysis.

#8-Post-trial onset discussion of the primary and secondary analyses

During 1998 and 1999 (and submitted in final form on April 30, 1999) the sponsor and the agency discussed the Statistical Analysis Plan. The Plan clarified the statistical test (and accompanying assumptions) to be used, and use of the logistic regression model was understandable, given the desire to be able to adjust for covariates. The model would be anchored at the 12 month point (the conventional approach), so it required imputation (LOCF) for dealing with withdrawals (see section 3.1.2.3.3 below).

The Statistical Analysis Plan also offered a modified population for the primary analysis, and a modified definition of "responder." The agency welcomed any additional analyses but noted that the protocol-specified analysis on all randomized patients should remain primary in order to maintain the prospective character of the experiment and to abide by the methodological

dictates of the null hypothesis paradigm, and noted that robust trends at the primary analysis may be supported by important secondary endpoints and analyses.

The specific proposals in the Statistical Analysis Plan were:

(1) A modified endpoint definition, which used a different cutpoint for the four indices.
(2) A modified population for the primary analysis – those treated for at least two months, rather than the all-randomized cohort.
(3) An additional covariate (SLEDAI>2, yes/no) in the model.
(4) A new endpoint defined as flare (NDA vol 1.44, p81):
new/worse CNS lupus, vasculitis, or myositis scored on SLEDAI & not present on previous visit
hem: plt<60K or Hb<7mg/dl or a fall of at least 3mg/dl
renal: proteinuria w/ pyuria or hematuria rx with new/inc. steroids or cytotoxics
steroids: inc. of at lease 2.5mg for at least 7 days for lupus
cytotoxics/anti-malarials: new/inc. use for at least 7 days for lupus
hospitalization for new lupus manifestation
other: "describe"

3.1.2.3 Protocol

3.1.2.3.1 Population, procedures: Inclusion criteria were (1) female sex, (2) diagnosis of lupus by criteria (1982 American College of Rheumatology criteria, see protocol appendix), (3) baseline SLAM score of at least 7, (4) steroid use of 10mg/d or less unchanged for the prior six weeks, (5) stable doses of cytotoxic drugs (azathioprine, methotrexate) or hydroxychloroquine for the prior six weeks, and (6) negative pregnancy test plus reliable, non-hormonal based, birth control if of child-bearing potential.

Reviewer's Comment: At the time of the design of this study, the published literature suggested a correlation between the SLAM and the SLEDAI, leading to the decision to use a SLAM>7 as a screen for patients with adequate disease activity. Later, after the results of trial GL94-01were interpreted as showing a better response in the SLEDAI>2 subset, and an analysis of the baseline SLEDAI and SLAM scores in patients enrolled to date in trial GL95-02 showed a weaker than expected correlation, the study was amended (see section 3.1.2.3.4, below) to include the SLEDAI>2 as an inclusion criterion and to expand enrollment for an additional 50 patients.

Exclusion criteria were as follows:

(1) history of breast cancer or cancer of the reproductive tract organs
(2) hemodialysis dependent

(3) treatment with ACTH, androgens, immunoglobulins, cyclophosphamide, cyclosporin A, or other immunosuppressive agents except azathioprine and methotrexate use with the last three months
(4) hypersensitivity to DHEA
(5) prior DHEA use within three months, or participation any prior DHEA study
(6) recent investigational agent use (within 30 days or 10 half-lives)
(7) any condition likely to prevent adequate compliance
(8) any serious EKG abnormality
(9) pregnancy or breast feeding.

Study plan consisted of screening and qualifying visits within 10 days, randomization, then four return visits for assessments at weeks 13, 26, 39, and 52, and a 58-week, off-therapy, visit. Patients discontinuing study medication prematurely were encouraged to be followed for 52 weeks.

3.1.2.3.2 Endpoints: The endpoint was defined as a by-patient responder test, consisting of improvement or stabilization of all four primary clinical variables, SLEDAI, SLAM, KFSS, PG, plus the absence of any "clinical deterioration". Clinical deterioration was defined as any of the following conditions, considered to reflect serious drug toxicity or worsening disease or both:

New onset diabetes
New gastric or duodenal ulcer not due to H. pylori requiring hospitalization or transfusion
New onset hypertension requiring drug therapy for at least three months
New myocardial infarction (by EKG or enzymatic criteria)
New steroid myopathy
New transaminase elevation (8-fold on one measurement, 3-fold on repeated measure)
New osteoporotic fracture
CNS: CVA, transverse myelitis, retinal vascular occlusion, new onset psychosis, new onset seizures refractory to therapy
Renal: New onset renal failure or progression to dialysis for at least three months
Pulmonary disease: New/worsened pulmonary hypertension, interstitial lung disease
Cardiovaascular: Refractory pericarditis, cardiomyopathy with hemodynamic compromise or refractory arrhythmia
Gastrointestinal: Ischemic bowel disease requiring resection
Vascular: Vasculitis resulting in infarction
Hematologic: Thrombocytopenia resulting in clinically significant hemorrhage with sequalae, persistent leukopenia (<1500) resulting in recurrent infections for at least three months
Medications: Any increase in concomitant methotrexate or azathioprine or beginning any new cytotoxic therapy during or 6 weeks

post-discontinuation, or any prednisone increase beyond limits specified in the protocol

3.1.2.3.3 Statistical considerations

Reviewer's Comment: As noted above, the log regression technique was specified in the protocol, although exactly how missing data would be handled was not explicit. It was always a goal to make the analysis as robust as possible to assertions that an effect seen was likely a treatment effect and not an effect due to differential withdrawal patterns, a problem that has plagued the interpretation of many prior rheumatic disease randomized trials. Eventually the imputation needed for the log regression analysis of the primary endpoint was determined to be the usual last-observation-carried-forward (LOCF).

The agency was also be interested in a simple way to compare durable responses, such as defining "success" as responding and completing 12 months of therapy (similar to the first primary endpoint in Trial GL94-01). (This was done – see Section 3.1.2.4.3.5, below). This approach would also have the advantage of being imputation free and thus possibly more valid, but it had the disadvantage of not being as powerful as the log regression analysis.

PRIMARY ANALYSIS (ORIGINAL PROTOCOL): The primary analysis was a comparison of the proportion of responders in all patients randomized by ITT, analyzed by logistic regression with treatment and center as factors. Covariates attaining 0.05 significance level for association with treatment assignment were included in the model, with eight covariates specified in the protocol (race, cytotoxic use, prednisone use, menopausal status, and baseline SLEDAI/SLAM /KFSS /PG). Score changes were calculated for each patient by subtracting the mean of the two baseline visits (screening and qualifying visits) from the mean of the four on-treatment visits (weeks 13, 26, 39, and 52 visits).

SECONDARY ANALYSES

(1) An all-randomized cohort was to be analyzed by comparing the mean changes for SLEDAI, SLAM, KFSS, PG, physicians global (also by 10cm VAS), SF-36, using the difference between the two baseline observations and the on-treatment observations (e.g. weeks 13, 26, 39, and 52), in a two-way analysis of covariance model with treatment and center as factors, and baseline as a covariate. Treatment-by-baseline and treatment-by-center interactions were included in the model.
(2) A trial-completing cohort (on or off study drug) was to be analyzed for all the same measures as above except the SF-36, by the same method.
(3) A time-to-discontinuation of study drug from any cause would be analyzed by means of a Cox regression model with treatment as a factor.
(4) A time-to-clinical-deterioration was to be analyzed using Kaplan-Meier curves and analyzed by the method of Cox regression using the same eight covariates as in the primary analysis.
(5) DEXA scan summaries (sub-study of bone density)

ALTERATIONS TO THE ANALYSIS PLAN MADE BY THE STATISTICAL ANALYSIS PLAN & OTHER MATERIAL DISCUSSED THROUGHOUT 1998 AND SUBMITTED IN FINAL FORM ON APRIL 30, 1999

See **Section 3.1.2.2** for further discussion

(1) The population for the primary efficacy responder analysis was redefined as all patients who have been on-treatment for at least 2 months.
(2) A responder analysis would be done on the SLEDAI>2 subset.
(3) Center was eliminated as a factor in the logistic regression, because "there are 23 investigation sites with few patients in many study centers."
(4) Three covariates – SLEDAI>2 (yes/no), prednisone use, and menopausal status – were now specified, and they were designated as "clinically important covariates" to be included in the logistic regression model used for the adjusted analysis, along with "any (other) imbalanced baseline covariate which attained a 0.05 significance level for association with treatment asignment." This adjusted analysis was described as "to confirm the result of the primary (unadjusted) analysis". The concern about patients with baseline SLEDAI>2 was the object of the amendment (July, 1997) restricting future enrollment to patients with a SLEDAI>2, in light of the results of trial GP94-01 (see above). The other two clinically important covariates were baseline prednisone (yes/no) and menopausal status (pre/other). A stepwise logistic regression model was used to identify the "best " model. The data showed only the p-values for "treatment effect" and "treatment by baseline SLEDAI interaction" were significant (0.0217 and 0.0010, respectively). However, the analysis was conducted using all three covariates.
(5) The concept of a "flare" was defined (see **Section 3.1.2.2**), and a time-to-flare analysis proposed.
(6) Comparisons using a different definition of responder in the following:
    all-randomized cohort
    2-month treatment subset
    SLEDAI>2 subset
    2-month treatment / SLEDAI>2 sub-subset.

    Reviewer's comment: "Two-month treatment" cohort proposal:

    As noted in Section 3.1.2.2 above, it is here that the sponsor first proposes the "two-month treatment cohort" for the primary analysis. These are patients on assigned drug for more than 60 days, having post-60 day data, and without major protocol violation. Analysis of a subset of the randomized cohort always has the potential to introduce bias, which may be difficult to predict, identify, or control for.

SAMPLE SIZE DETERMINATION: Given that there were no prior data with which to use to power this study, no sample size calculation could be done. The original sample size of 300 randomized patients was mainly based on feasibility. The protocol was amended on August 4, 1997 to increase the sample size (see **Section 3.1.2.3.4**, below).

HANDLING OF WITHDRAWALS: The final protocol (January 31, 1996) did not fully detail the handling of withdrawals and other sources of missing data. These details were addressed in a teleconference of June 30, 1999 and a subsequent memo sent by the sponsor on July 20, 1999 (appended to the protocol, attached). The eventual statistical tests to be employed for the primary and secondary analyses were specified in the Statistical Analysis Plan submitted April 30, 1999. These invoked the usual LOCF imputation technique.

3.1.2.3.4: Protocol amendments

Amendment #1 (August 4, 1997): This amendment provided for an additional 50 patients with a baseline SLEDAI score of >2 as an inclusion criterion because a subset analysis of trial GP94-01 showed a greater point estimate difference in this subset compared with patients with baseline SLEDAI scores of two or less.

Amendment #2 (April 21, 1999): This amendment provided for uterine and breast monitoring of post-menopausal patients in Trial GL95-02, in light of the finding from Trial GL94-01 of elevated hormone levels in those post-menopausal patients not on hormone replacement therapy with the use of DHEA.

3.1.2.3.5: Additional filings, not as official protocol amendments

As noted above (**Section 3.1.2.3.3**) the sponsor submitted a finalized Statistical Analysis Plan on April 30, 1999.

3.1.2.4: Results

3.1.2.4.1: Patient accountability

	DHEA	PLC	TOTAL
Total randomized	189	192	381
Completed on assigned agent	124	142	266
Early Termination of Drug	65	50	115
Reason:			
Inefficacy	11	9	20
Adverse event	31	18	49
Other	23	23	46

CHARACTERISTICS OF PATIENTS WITHDRAWING FROM THE TRIAL

Reviewer's Comment: An attempt to understand missing data is central to trying to understand trial results, whether positive or negative, and for this reason, information regarding withdrawals is provided here. These

data will supplement the efficacy analyses supplied below (Section 3.1.2.4.3)

Baseline comparisons of withdrawals with the full-randomized cohorts are shown in Table 19 (appendix).

No significant differences are apparent here.

Survival plots for all withdrawals, inefficacy withdrawals, and adverse event withdrawals are shown in Table 20 (3 sheets, appendix). A trend (p=0.077) towards more DHEA discontinuations for all reasons, driven mainly by those for adverse events (p=0.039) was seen. A list of all patients terminated early because of an adverse event is given in Table 21 (appendix). Many of these patient adverse events may reflect androgenic and estrogenic effects of DHEA.

An analysis of the prednisone dosages of the withdrawals showed that a large majority had no change in their prednisone dose during their time on trial. Specifically, of the 65 DHEA withdrawals, 54 had no change in prednisone, and of the 50 PLC withdrawals, 40 had no change. Nearly half of these (24/54 for DHEA, 18/40 for PLC) entered and exited the trial on no steroids at all. Of withdrawals who did show a change in steroid dose during the trial, the 11 patients on DHEA showed the following range of changes:

-10, -2.5, -1, 2, 2, 3.75, 5, 5, 5, 9, and 12mg,

whereas, the 10 withdrawals on PLC showed the following range:

-5, -0.5, -1, 2.5, 5, 5, 10, 10, 50, and 57.5mg.

The last two patients (#14483-withdrew on day 265, and #36515-withdrew on day 77), were the only ones exiting for treatment of a major lupus flare with high dose steroids. With the possible exception of these two placebo patients, no large difference across the two arms is seen.

Finally, 14/65 (22%) of the DHEA withdrawals were responders, compared to 58/189 (30%) of the entire DHEA cohort. 7/50 (14%) of PLC withdrawals were responders, compared with 52/192 (27%) of the entire cohort.

3.1.2.4.2 Patient comparability at entry

DEMOGRAPHICS

	DHEA	PLC
Age (mean, range)	44.4 (19-69)	43.8 (18-68)
Race*(%white/black/other)	77/12/11	71/17/12
Pre-menopausal*	56%	54%
Cytotoxic use*	17%	15%
Prednisone (mean, mg/d)	3.5mg	3.7mg
Pre-existent hypertension	35%	31%
Current smoker	19%	14%

BASELINE STATUS

SLEDAI* (range)	6.6 (0-18)	5.8 (0-24)
SLAM* (range)	12.2 (6.5-21)	12.0 (4-21)
KFSS*	5.5	5.6
PG* (10cm VAS)	55.2	55.4
IG (10cm VAS)	30.2	30.3
SF-36 physical	31.1	31.6
SF-36 mental	42.5	41.7
SLICC damage index (range)	1.3 (0-7.0)	1.3 (0-9.0)

* these eight factors were protocol-specified covariates (see above)

Enrollees by baseline SLEDAI are shown graphically in Table 22 (2 sheets, appendix)

3.1.2.4.3 Efficacy endpoint results

Reviewer's Comment:

The protocol was submitted on January 31, 1996. Later amendments contained proposals for a modified primary endpoint and a modified primary population. The Results Section below has been arranged to display the original primary and secondary analyses followed by the modified analysis plans submitted subsequently by the sponsor.

3.1.2.4.3.1: Primary analysis (original protocol)

The primary analysis was a comparison of responders in the all-randomized cohort, by logistic regression using treatment and centers as factors, and with whichever of eight covariates attained a 0.05 significant association with treatment included.

	DHEA	PLC
Responder	58/189 (30%)	52/192 (27%)
Non-responder	131/189	140/192

P=0.436 (adjusted per initial protocol)
p=0.486 (adjusted per Statistical Analysis Plan)
p=0.438 (unadjusted)

3.1.2.4.3.2: Secondary analyses

Secondary Analysis #1: Comparison in the all-randomized cohort of means of baseline and change from by baseline by ITT in SLEDAI, SLAM, KFSS, PG, plus the investigator global (IG) and the two components of the SF-36, the physical component (SF-36PC) and the mental component (SF-36MC), using a two-way analysis of variance model with treatment (ANCOVA).

		SLEDAI	SLAM	KFSS	PG	IG	SF-36PC	SF-36MC
DHEA	baseline	6.5	12.2	5.5	55.2			
	change from baseline	-2.2	-2.9	-0.3	-5.1	-6.2	2.0	2.3
PLC	baseline	5.8	12.0	5.6	55.4			
	change from baseline	-1.8	-2.9	-0.3	-4.7	-6.4	1.8	2.3
P VALUE		0.25	0.38	0.45	0.86	0.93	0.79	0.96

Scatterplots for baseline and final prednisone doses for all patients by treatment arms are shown in Table 23 (appendix).

Secondary Analysis #2: Comparison of trial-completing (on or off assigned treatment) of mean changes in SLEDAI, SLAM, KFSS, PG, physician's global, and prescribed prednisone dosage using a two-way analysis of variance model with treatment.

Result: Not submitted.

Secondary Analysis #3: Time-to-discontinuation by means of a Cox regression model with treatment as a factor.

Result is shown above in **Section 3.1.2.4.1**.

3.1.2.4.3.3 Pre-specified bone density sub-study

This sub-study was done at 8 of the 23 centers, and only on patients who had been on prednisone for at least 6 months. Thirty-seven (37) patients were enrolled, 18 on DHEA and 19 on PLC. The results showed an increase in bone density in DHEA patients compared to a decrease in the PLC group. There was a 2.08% increase in bone density measured at the hip in the DHEA treated patients, compared to a –0.16% loss in the PLC treated patients (p=0.080), and a 1.83% versus –1.78% in the lumbar spine (p=0.004).

> ***Reviewer's Comment: This sub-study was encouraged because at the beginning of trial GP95-02 there remained the question whether DHEA acted by simply making prednisone more bioavailable or augmenting its pharmacodynamic effect, resulting in an osteoporotic effect. It was recognized that success in this aspect would make an important contribution to any risk / benefit assessmemt of chronic DHEA use versus chronic corticosteroid use.***

3.1.2.4.3.4 Subset analyses proposed in amendments and in April 20, 1999 Statistical Analysis Plan:

Demographics of subgroups populations

Two-month treatment subset n=346
SLEDAI>2 subset n=293
Two-month treatment AND SLEDAI>2 subset n=265

	DHEA	PLC	TOTAL
Total randomized population	189	192	381
Total two-month treatment population	170	176	346
Total excluded	19	16	35
Reason for excluding			
Missing post-baseline data (all 4 variables)	19	13	32
With data but treatment < 60d		2	2
Major protocol violation	1		1
Completed on assigned agent	124	142	266
Early Termination of Drug	46	35	81
Reason:			
Inefficacy	10	9	19
Adverse event	19	8	27
Other	17	27	44

Two-month treatment subset: baseline status

	mean (median)	
	DHEA	PLC
SLEDAI (0-105)	6.5 (6)	5.9 (5)
SLAM (0-60)	12.3 (12)	12.0 (12)
KFSS (0-7)	5.5 (5.9)	5.6 (5.7)
PG (0-100)	55.2 (57)	55.1 (57)
SF-36 physical (0-100)	31.5 (30.9)	31.3 (30.2)
SF-36 mental (0-100)	42.3 (43.1)	42.1 (42.3)

RESULTS: TWO-MONTH TREATMENT SUBSET

1. Responder comparison, unadjusted and adjusted logistic regression (with SLEDAI, prednisone use, and menopausal status as covariates).

	DHEA		PLC
Responder	**60/170**		**52/176**
Non-responder	**110/170**		**124/176**
		p=0.311 (adj.)	
		p=0.254 (unadj.)	

2. Comparison of mean change from baseline by ITT in SLEDAI, SLAM, KFSS, PG, physician's global, SF-36, and prescribed prednisone dosage, using a two-way analysis of variance model with treatment.

	DHEA	PLC
SLEDAI	**-2.24**	**-1.72**
SLAM	**-3.10**	**-2.65**
KFSS	**-0.33**	**-0.39**
PG	**-6.24**	**-4.35**
Physician Global	**-5.64**	**-5.19**
SF-36-mental	**2.64**	**1.80**
SF-36-physical	**1.76**	**1.71**

No comparison reached statistical significance at 0.05

RESULTS: SLEDAI>2 SUBSET

1. Responder comparison by ITT, by unadjusted and adjusted logistic regression (with SLEDAI, prednisone use, and menopausal status as covariates).

	DHEA		PLC
SLEDAI>2 subset	**55/147**		**42/146**
		p=0.117	
		(0.127adj)	

2. Comparison of mean change from baseline by ITT in SLEDAI, SLAM, KFSS, and PG, using a two-way analysis of variance.

	DHEA	PLC	p value
SLEDAI	**-3.2**	**-2.5**	**0.146**
SLAM	**-3.2**	**-2.7**	**0.188**
KFSS	**-0.3**	**-0.3**	**0.610**
PG	**-7.2**	**-3.0**	**0.062**

RESULTS USING REVISED RESPONDER DEFINITION

This modified definition used a slight deterioration as the cutoff for each measure, specifically, 0.5 for the SLEDAI, 1.0 for the SLAM, 0.5 for the KFSS, and 10 for the PG. The range of the measures are 0-105 (but usually <45) for the SLEDAI, 0-86 for the SLAM, 0-7 for the KFSS, and 0-100 for the PG.

Reviewer's Comment: While cutoffs are needed for each measure in the responder definition, there is no clear rationale to favor the specific cutoffs provided in the modified responder definition. Measurement variability is always present, and may well have been expected to be prominent in some of these measures, but that variability does not determine what cutoff to use. There had been discussion of this issue in the protocol design stage (see Section 3.1.2.2). The Statistical Review explores this decision more deeply.

Responder comparison by ITT, by unadjusted and adjusted logistic regression, using the revised definition (right hand column below), shown along side the results using the protocol definition of responder (left hand column).

Responder definition	protocol defined		revised by amendment	
	DHEA	PLC	DHEA	PLC
all-randomized cohort	58/189	52/192 p=0.438 (0.486adj)	97/189	81/192 p=0.074 (0.086adj)
SLEDAI>2 subset	55/147	42/146 p=0.117 (0.127adj)	86/147	65/146 p=0.017 (0.020adj)
two-month treatment subset	60/170	52/176 p=0.254 (0.311adj)	99/170	80/176 p=0.018 (0.026adj)
two-month & SLEDAI>2 treatment subset	56/132 p=0.068	42/133 (0.082adj)	87/132 p=0.005	65/133 (0.008adj)

Reviewer's Comment: These p-values are "nominal"; statistical interpretation is problematic.

8.1.2.4.3.5 Other exploratory analyses

A. PROPORTION OF RESPONDERS WHO COMPLETED TRIAL: This analysis is akin to that used in the design of GL94-01, where a responder needed not only to meet certain clinical measurement criteria but to also have finished the trial. This approach has the advantage of being free of any imputation, such as ITT/LOCF. Here, success would be defined as a protocol responder who also completed the twelve-month trial on assigned therapy, and all others would be categorized as failures. The statistical test would then be a simple Fisher's exact. The results here showed:

DHEA	44/189
PLC	45/192

p=1.000

Reviewer's Comment: An analysis like this has intuitive clinical appeal for assessing a chronic condition because it captures not only efficacy but durability. On the other hand, it may be less sensitive to treatment effects.

B. TIME TO FIRST FLARE: These analyses are done using the "flare" definition supplied in the Statistical Analysis Plan (see **Section 3.1.2.2**). These analyses were first done on three cohorts: the full cohort of randomized patient (n=381), the 2-month

treatment subset (n=346 total), and 2-month treatment/SLEDAI>2 subset (n=265 total). Two caveats of the analysis are important here: (1) the "window" for capturing the flare was on assigned treatment plus an additional seven days thereafter, and (2) no flare occurring before the 60 day point was counted. Using the log rank test, the following were found, comparing DHEA with PLC:

1) full cohort of randomized patients
   DHEA 45/189 (24%)
   PLC 57/192 (30%)

2) two-month treatment subset
   DHEA 37/170 (22%)
   PLC 47/176 (27%)
   p=0.335

3) 2-month treatment/SLEDAI>2 subset
   DHEA 31/132 (23%)
   PLC 41/133 (30%)
   p=0.201

Reviewer's Comment: A "flare" analysis also has intuitive clinical appeal. Although it would have been preferable for the flare definition to have been prespecified before the trial began, the flare definition used here was derived from another database not available at the outset of Trial 95-02. The seven-day window is too short, but this could not be remedied post hoc.

C. TIME TO DETERIORATION: The protocol specified a responder needed to have not worsened in the four efficacy measures (SLEDAI, SLAM, KFSS, and PG), ***plus*** not to have demonstrated any of the protocol specified "major clinical deterioration" events (see **Section 3.1.2.3.2**). Consequently, it was of interest how many patients demonstrated this "major deterioration". The analysis showed:

DHEA patients with major deterioration	16/189
PLC patients with major deterioration	16/192

p=0.869 (log rank)

This is shown graphically in Table 24 (appendix)

If, in addition, one adds "new or worsening hypertension" (see **Section 3.1.2.4.4.2**, below) to the list of clinical events qualifying as a "major deterioration," the results differ little:

DHEA patients with (modified) deterioration	18/189
PLC patients with (modified) deterioration	18/192

Only two patients in each arm would be subtracted from the responder list if new or worsening hypertension is added.

D. PROTOCOL RESPONDER BY BASELINE SLEDAI: In light of the results of trial GL94-01, an analysis of response as a function of the baseline SLEDAI would be of interest. What follows is a table of the proportion of responders as a function of baseline SLEDAI.

All randomized cohort	DHEA	PLC
baseline SLEDAI		
0-2	3/42 (7%)	10/46 (22%)
2-4	9/29 (31%)	11/40 (28%)
4-8	30/65 (46%)	24/68 (35%)
>8	16/53 (30%)	7/38 (18%)
2-month treatment cohort		
0-2	4/38 (11%)	10/43 (23%)
2-4	9/25 (36%)	11/36 (31%)
4-8	30/58 (52%)	24/62 (39%)
>8	17/49 (35%)	7/35 (20%)

E. PROPORTION OF PATIENTS WHO FLARE BY BASELINE SLEDAI

All randomized cohort	DHEA	PLC
baseline SLEDAI		
0-2	9/42 (21%)	7/46 (15%)
2-4	3/29 (10%)	12/40 (30%)
4-8	14/65 (22%)	18/68 (27%)
>8	19/53 (36%)	20/53 (38%)
2-month treatment cohort		
0-2	6/38 (16%)	6/43 (14%)
2-4	1/25 (4%)	10/36 (28%)
4-8	13/58 (22%)	13/62 (21%)
>8	17/49 (35%)	18/35 (51%)

F. DESCRIPTIVE STATISTICS FOR SECONDARY VARIABLES BY BASELIINE SLEDAI

Below are the mean changes for the variables (seven) specified for secondary analysis, plus the SLICC variable, sorted by baseline SLEDAI subgroup. They are included for data analysis by those interested in lupus measurement. There are no obvious clinical interpretations of these data.

All randomized cohort

	----------DHEA----------				----------PLC----------			
baseline SLEDAI →	0-2	2-4	4-8	>8	0-2	2-4	4-8	>8
SLEDAI	1.00	-0.84	-2.51	-5.14	0.86	-0.53	-2.24	-5.06
SLAM	-2.89	-3.41	-3.36	-2.80	-2.68	-2.55	-3.07	-2.01
KFSS	-0.33	-0.27	-0.44	-0.22	-0.74	-0.51	-0.29	0.01
PG	-2.85	-5.72	-8.83	-6.12	-9.21	-4.44	-5.49	3.08
Inv. Global	-6.51	-4.56	-8.83	-6.03	-7.27	-3.59	-4.09	-5.51
SF-36-physical	1.39	0.95	2.61	1.45	4.11	-0.28	2.03	-0.03
SF-36-mental	3.69	1.32	3.36	1.62	2.41	3.15	1.91	-0.57
SLICC	-0.03	-0.11	-0.04	-0.17	-0.03	-0.07	0.04	-0.21

2-month treatment cohort

	----------DHEA----------				----------PLC----------			
baseline SLEDAI→	0-2	2-4	4-8	>8	0-2	2-4	4-8	>8
SLEDAI	1.00	-0.84	-2.51	-5.14	0.90	-0.53	-2.31	-5.21
SLAM	-2.89	-3.41	-3.36	-2.80	-2.69	-2.55	-3.00	-2.07
KFSS	-0.33	-0.27	-0.44	-0.22	0.76	-0.51	-0.28	-0.01
PG	-2.85	-5.72	-8.83	-6.12	-8.98	-4.44	-5.35	3.41
Inv. Global	-6.51	-4.56	-5.19	-6.03	-7.25	-3.59	-4.36	-5.77
SF-36-physical	1.39	0.95	2.61	1.45	4.19	-0.28	2.05	0.07
SF-36-mental	3.69	1.32	3.36	1.62	2.31	3.15	1.88	-0.46
SLICC	-0.03	-0.11	-0.04	-0.17	-0.03	-0.07	0.04	-0.22

G. PROTOCOL RESPONDERS BY RACE: Number of patients too small for analysis.

H. PROTOCOL RESPONDERS BY MENOPAUSAL STATUS: No differences found.

I. PROTOCOL RESPONDERS BY BASELINE PREDNISONE DOSE: In dividing patients into four categories: 0mg/d, >0-5mg/d, >5-7.5mg/d, >7.5-10mg/d, no differences were found.

J. PROTOCOL RESPONDERS AS A FUNCTION OF A LABORATORY PARAMETER PRESUMED RELATED TO THE TREATMENT HYPOTHESIS: These studies were done for study GL94-01 (see **Section 3.1.1.4.3.4**) in the form of scatter plots, but they were not done for study GL95-02.

3.1.2.4.4 Safety comparisons

3.1.2.4.4.1 Extent of exposure

	DHEA200mg	PLC
N	189	192
Mean (day)	288	308
Medium (day)	359	362
SD (day)	119	107

3.1.2.4.4.2 Adverse events

ADVERSE EVENTS OCCURRING IN AT LEAST 10% OF PATIENTS: Table 25 (appendix) shows adverse events occurring in at least 10% of patients, sorted by the occurrence rate in the DHEA patients. Acne, myalgia, hirsutism, and stomatitis ulcer were more common in the DHEA cohorts.

Of adverse events occurring in <10% of patients, the following were reported significantly more frequently in the DHEA group

	DHEA	PLC
hot flashes	7 (3.7%)	0
creatinine increase	5 (2.6%)	0
hematuria	7 (3.7%)	0
back pain	15 (7.9%)	6 (3.1%)

ADVERSE EVENTS OCCURRING IN AT LEAST 5% OF PATIENTS: Table 26 (appendix) shows adverse events possibly or probably attributed to the drug for events occurring in at least 5% of either group.

SERIOUS ADVERSE EVENTS REGARDLESS OF RATE OR ATTRIBUTION: Table 27 (2 sheets, appendix) shows adverse events assessed as severe.

POST-HOC ANALYSIS OF ADVERSE EVENTS BY SLEDAI SUBSETS: This analysis is the same as was done for trial GL94-01; it proved unrevealing. Full data are in the appendix (Table 28, 17 sheets).

DEATHS: There were five deaths during the study and follow-up period. All were on PLC. The causes of death were one with pulmonary hypertension and restrictive lung disease, one with non-Hodgkin's lymphoma, one by sudden death, and two by suicide.

MALIGNANCY: Three patients on placebo were diagnosed with cancer during trial GL95-02. Two patients on DHEA in the one-year, open-label extension

study GL95-01, and one at follow-up after GL95-01, were also diagnosed with cancer, these three all being breast cancer.

PLC#1	carcinoma of the breast
PLC#2	carcinoma of lung
PLC#3	non-Hodgkin's lymphoma
DHEA#1	infiltrating ductal cancer of breast
DHEA#2	infiltrating ductal cancer of breast
DHEA#3	breast cancer (NOS)

The three patients on DHEA were:

Patient #12013 (also called #12578): A 49 year-old, proportedly still having menstrual periods, who had received 12 months DHEA at 200mg/day in GL95-02, and another 12 months DHEA at 200mg/day in GL95-01, then subsequently self medicated with a dietary supplement of DHEA, of unknown strength, for an additional 4 months until the time of diagnosis. The sponsor has said that this patient was "still having menstrual periods."

Patient # 45006 (also called #45588): A 61 year-old who had received a total of 15 months of DHEA 200mg/d (12 months in GL95-02, followed by 3 months in GL95-01) before the time of diagnosis.

Patient #18035 was a 49 year-old, still having menstrual periods, who had been in both trials GL95-02 and GL95-01, totally 24 months of DHEA at 200mg/day. It was not reported whether she used supplemental DHEA following, but four months after completing GL95-01 she presented with pain in the breast, four months later swelling, redness and a discharge appeared. A mammogram was done two months later, read as probably benign. After an additional four months she presented with breast bleeding for three months and was hospitalized, had cancer diagnosed (now a total of 13 months after the end of GL95-01, and died one month later. The menopausal status of this patient has not yet been reported.

> ***Reviewer's Comment: Although it is possible that a signal is occurring here, it is overwhelmingly more likely due to chance. Under almost all circumstances involving carcinogenicity, even large controlled studies, and even larger observational databases rarely are able to supply confident risk estimates.***

HYPERTENSION: The following is a table of patients showing new onset or worsening hypertension during this trial:

	DHEA	PLC
new onset HT	8	8
increased HT	9	15
increased rx needed	12	12

Of those patients with a prior history of, or pre-existent, hypertension, 7 of 35 on DHEA developed hypertension, compared to 8 of 23 on PLC.

Blood pressure was further analyzed, showing the following.

	Mean (median) values	
	DHEA200mg	PLC
DBP-Initial	76.8 (78)	77.0 (75.5)
Change	-0.1 (0.0)	-2.5 (-2.0)
SBP-Initial	122.1 (120)	121.8 (120)
Change	0.3 (0.0)	-2.9 (-4.0)

WEIGHT: There was no meaningful difference in weight over the trial in any arm.

PREGNANCY: Despite precautions in the protocol, two patients on DHEA (one for 30 days, the other for 74 days) became pregnant. Both were reported as resulting in normal infants.

MAMMOGRAPHY / UTERINE ULTRASOUND: As mention above (**Section 8.1.1.4.4.3.1**), a monitoring program was instituted to access on treatment post-menopausal women in GL95-02 and GL95-01 because of concern of high estradiol levels attending DHEA therapy in some of the post-menopausal women in trial GL94-01. The timing of this amendment was such that it occurred at nearly the end of GL95-02, so that there could have been mammography or uterine ultrasound or biopsy data on GL95-02 patients prior to the amendment which would not have been captured by the sponsor. Unless these data were sent to the sponsor unprompted, they could be obtained only by retrospective chart review.

How many patients enrolled in the open label trial GL95-01 were post-menopausal / off hormone-replacement-therapy is unknown. The Four-Month Safety Update (February 9, 2001) describes 43 DHEA patients “available for study” by mammography under this amendment. Of these, 29 had a total DHEA of year or less, and 14 had a DHEA exposure of between one and two years.

There is no DHEA sponsor data beyond the 24-month point except for one center (Johns Hopkins) where patients could remain on DHEA for two years. One underwent breast biopsy revealing a degenerating fibroadenoma.

Uterine ultrasound (transvaginal) has been done on 24 patients under this amendment. Of these 16 have a DHEA exposure of a year or less, and 8 an exposure of between one and two years. Eight were abnormal, of which three were biopsied, but none were reported as showing hyperplasia or neoplasia.

Reviewer's Comment: The on-treatment and follow-up database is too far too small for any conclusions regarding whether there is a DHEA effect mediated via estrogen, testosterone, or some other route.

3.1.2.4.4.3 Laboratory Data

3.1.2.4.4.3.1 Descriptive statistics

ESTRADIOL: There was an increase in estradiol associated with DHEA treatment levels in those post-menopausal patients not on exogenous estrogens/progestins (see Table 29, appendix). As noted earlier (Section 8.1.2.4.4.2) this finding in GL94-01 led to the program of monitoring breast and uterine tissue.

FSH / LH: A mean change was seen for FSH of –0.7 from a baseline of 37.0 on DHEA, compared to 1.8 and 33.4 on PLC. LH showed a change of 1.9 from 20.0 on DHEA, compared to –5.2 and 17.3 on PLC.

TESTOSTERONE: Testosterone levels also showed the expected, increasing on DHEA (Table 30, appendix), without a difference between pre- and post-menopausal groups.

LIPID LEVELS: The values for total, HDL-, LDL-cholesterol, the cholesterol/HDL-cholesterol ratio, and triglycerides are shown in Table 31 (appendix), with many between-group comparisons reaching "statistical significance".

Reviewer's Comment: The lipid profile changes noted in this study need to be considered in a risk/benefit analysis.

COMPLEMENT: The values for complement levels, C3 and C4, are also shown in Table 31 (appendix), with the changes in C3 reaching "statistical significance".

OTHER LABORATORY MEASURES: Mean 24 hour urinary protein was reported in the text of the NDA as showing an increase in both groups: 171mg for DHEA versus 48mg for PLC. A dsDNA titer rise of 20 was seen in patients on

DHEA, compared to 6 in PLC, and urinary RBC/hpf of –0.6 in DHEA, compared to 0.4 in PLC. These data were all subsequently analyzed as below.

3.1.2.4.4.3.2 Analysis by reviewer of differences in parameters suggesting a worsening of renal disease and/or worsening serology on DHEA. This is identical to the analysis done for trial GL94-01 – see 3.1.1.4.4.3.2).

Definition for worsening were the following, all requiring a second, follow-up confirmatory value, unless after one abnormality the patient withdrew specifically for worsening of one of these parameters. In this trial, as in GL94-01 no patient qualified as a case in this analysis by this route.

Parameter	normal at baseline	abnormal at baseline
Urinary RBC count (nml: 0 RBC/hpf)	increase to $\geq$10	double to $\geq$10
24 hr urinary protein (nml $\leq$150mg/d)	double to >150mg/d	double
complement C3 (nml: 85-193)	25% decrease	25% decrease
complement C4 (nml: 12-36)	25% decrease	25% decrease
anti-double-stranded DNA (nml: 0-3.6)	double	double

Results:

	Patient number				
	urine RBC	urine protein	falling C3	falling C4	rising DNA
PLC	18717	15469	12580	12580	3609
	41439	15523	14483	14551	14483
	48819	15791	18747	20769	14551
		18617	23650	23725	15704
		18692	38605	35568	15793
		18719	46643	46643	20592
		18732	46696	46696	46696
		18768			
		20457			
		23727			
		28806			
		38402			
		38423			

	38736			
	43613			
	46643			
	48813			
	48819			
	49760			
	49779			
	49803			
DHEA 19454	12476	15521	15521	18724
19454	12577	15636	37493	18745
	14406	18720	38421	20684
	14550	20684		20770
	15414	20770		23485
	15524	36640		38401
	15702	38445		38445
	15794	38608		38478
	16557	48814		41535
	18722			45600
	18724			48814
	18731			
	18746			
	19454			
	21410			
	23649			
	32466			
	33441			
	33665			
	35654			
	36514			
	36570			
	38783			
	41535			
	41536			
	43601			
	43615			
	45563			
	45598			
	46641			
	46693			
	48816			
	48818			
	48820			

If one determines the number of patients, none counted twice, with at least one laboratory test suggesting new or worsening lupus disease, the following are obtained.

PLC	n=38
DHEA	n=51

The number of patients with more than one such suggestive test are:

PLC	n=5
DHEA	n=8

Reviewer's comment: Similar to study GL94-01, there were signals in GL95-02 suggesting decreasing complement and increasing proteinuria in the DHEA treated subjects. These result in the two trials constitute a "signal" that needs further exploration.

3.1.2.5 Conclusions

1. ***The results of the original protocol analysis show a p=0.436 for the primary endpoint, and p values ranging from 0.25 to 0.86 for the four secondary endpoints.***
2. ***Analyses in accordance with the sponsor's modified final analysis plan show show p values which reachnominally significant <0.05 values. The modified analysis plan included use of a subset of patients with baseline SLEDAI>2 and DHEA exposure of >60 days, and modification of the endpoint definition.***
3. ***Safety signals seen in study 94-01 related to the hormonal effects of DHEA as well as lipid profile changes were replicated in this study.***
4. ***A safety signal related to labs associated with possible renal toxicity seen in study 94-01 was replicated in this study.***

3.1.3 Study GL95-01: Open-label extension for patients from Studies GL94-01 and GL95-02 on 200mg/d DHEA.

This is an ongoing study of patients electing to continue DHEA after completion of blinded, controlled studies GL94-01 and GL95-02, to enable patients to receive up to a maximum of one year of DHEA. The patients were begun on 200mg/d of DHEA, but they could decrease for intolerance. Approximately 10-20% have elected to reduce their dose. No information was given in the NDA comparing those electing to enter this trial with those declining. At the time of the NDA submission, a total of 314 of approximately 370 patients enrolled to date are included in the interim analysis supplied in the NDA. (The subsequent Four-month Update, submitted on February 9, 2001, notes 572 patients were eligible, of which 371 elected to enroll.) 192 had been on DHEA and 179 had been on PLC in earlier trials. Patients were seen every three months and had the following

collected every six months: SLEDAI, patient and physician global, KFSS, and the short-form SF-36 health status survey.

Table 32 (appendix) lists adverse events occurring in at least 5% of the patients.

Adverse events ranked by severity and by race are noted below in Tables 33 and 34 (appendix).

Adverse events leading to dose reduction are shown in Table 35 (2 sheets, appendix).

A by-patient description of all those experiencing severe adverse events (including death in four patients) is given in Table 36 (16 sheets, appendix). Many of these events are common in the lupus population.

IMPORTANT NOTE:

There were three cases of breast cancer reported, one during and two after DHEA exposure in GL95-01. These are described in the report for GL95-02 above (**Section 3.1.2.4.4.2**), and they are further commented on in the Oncology and OPDRA Consults.

3.1.4 Trial GL97-01: RCT DHEA in men with lupus (ongoing).

This is a small RCT comparing DHEA 200mg/d with placebo in men with lupus. It is a one-year, double-blind study, followed by a one-year open-label extension. The intended size was 20 patients per arm, but recruitment has been slow because of the relative rarity of the disease in men. One patient was found to have prostate cancer shortly after his crossover from placebo to DHEA. As of the Four-month Update, nine patients have completed one year, and four have prematurely discontinued. Twenty-eight patients have been enrolled.

REVIEW APPENDIX – RESULTS OF OTHER TRIALS

1. RCT (van Vollenhoven #1): Comparison of 200mg DHEA and placebo in mild to moderate SLE (Arthritis Rheum 38: 1826-31, 1995)

DESIGN

A three month, double-blind, placebo-controlled RCT done by Dr. Van Vollenhoven at Stanford, compared DHEA 200mg/day and PLC in 28 patients with mild to moderate SLE. Minimal renal disease, defined as 1+ proteinuria, at most, and stable for six months, plus a normal creatinine, was an entry requirement. It is important to note that background NSAIDs, steroids, and hydroxychloroquine were permitted, including changes in these medications during the trial.

DEMOGRAPHICS

	DHEA	PLC
Number	14	14
Caucasian	9	11
African-American	5	1
Other	0	2
Age	33.9 yr.	38.5 yr.
Duration of disease	131 mo.	100 mo.
SLEDAI score	9.8	6.1
Average Prednisone Dose	15.8 mg	7.7 mg.

EFFICACY RESULTS

	Change from baseline		p value*	
	DHEA	PLC	unadj.	adj.
SLEDAI	-1.7	0.8	0.09	0.21
Patient global	-11.5	2.4	0.14	0.002
Investigator global	-3.1	1.1	0.47	0.32
Prednisone dose	-3.2	2.0	0.11	0.31

- ANOVA, adjusted for baseline, SLEDAI, prednisone, and globals

SAFETY RESULTS

Events seen, without regard to attribution, were as shown below. As will be seen, these reflect the common, and in many cases expected AEs seen in larger controlled databases (trials GL94-01 and GL95-02, below).

	DHEA (n=14)	PLC (n=14)
acne	8	3
alopecia	1	0
amenorrhea	0	1
aphasia	1	0
depression	3	0
hair disorder	0	1
hirsutism	2	4
menometrorrhagia	1	2

hernia	0	1
nervousness	2	0
pain	1	0
palpitatons	1	0
peripheral edema	1	0
rash	1	2
seborrheic dermatitis	0	1
tinnitus	1	0
weight gain	2	0

2. RCT (van Vollenhoven #2): Comparison of 200mg DHEA and PLC in severe SLE (Lupus 8:181-7, 1999)

This was a 6-month randomized, double-blind, PLC control trial in severe lupus. Its design in a number of ways paralleled the ideas that went into the designs of GL94-01 and GL95-02. Entry criteria for this pilot study required one of the three major lupus manifestations noted below, having either newly developed or having been present despite appropriate conventional therapy (including high dose steroids in all patients, and immunosuppressive therapy in most) for at least one month prior to enrolment:

(a) lupus nephritis -- 2 of the following: proteinuria of 3gm/d, or 2gm with nephrotic syndrome, RBC casts or 20 rbc/hpf, or a fall in creatinine clearance by 30% in the past three months
(b) hematological lupus -- autoimmune hematolytic anemia with Hct <20 associated with reticulocytosis >5%, positive Coombs test, or decreased haptoglobin, or thrombocytopenia of <50,000
(c) serositis -- pleurisy, pericarditis, or peritonitis, as documented by appropriate imaging.

The primary endpoint was intended to capture stabilization of the major lupus manifestation (i.e., nephritis, hematological lupus, or serositis – see above) at the six-month timepoint. Specifically, this endpoint was defined in the protocol as attaining the following without the addition of immunosuppressive therapy:

(a) lupus nephritis – proteinuria of <2gm/day or less than half of baseline, and no nephrotic syndrome, PLUS the absence of rbc casts AND <20rbc/hpf, AND creatinine clearance fall by <20% from baseline.
(b) hematological lupus – hematocrit of >25 and platelet count above 50,000 for the five-month and six-month assessment
(c) serositis – absence of the objective manifetations of serositis for at least one-month prior to the six-month assessment

All other patients were considered non-responders (including any with the addition of immunosuppressive therapy). There were numerous secondary endpoints including the

SLEDAI, SLAM, patient and investigator golbal assessment, prednisone dosage, the HAQ, and relevant laboratory parameters.

Twenty-one patients were enrolled including three male patients (all three were randomized to PLC). Two patients were excluded from the analysis (one did not return, and could not be located; a second died just after randomization of an arrhythmia in the setting of nephritis and nephrotic syndome, electrolyte disturbances, and polyserositis). Characteristics of the nineteen patients in the analysis are shown below.

	DHEA (n=9)	PLC (n=10)
Age (mean, sd)	35.4 (8.0)	39.1 (16.6)
Female (%)	9 (100%)	7 (70%)
Race		
Caucasian	3	4
Afro.-Amer.	1	2
Latin-Amer.	1	1
Asian-Amer.	2	2
Pacific Isl.	2	1
Major manifestation		
Nephritis	8	6
Hematologic	1	0
Serositis	0	4

During the trial a tapering regimen for prednisone was suggested but not mandated. Six-month data were available in 16, and the primary response was evaluable in 19. 7/9 DHEA patients showed a response, compared to 4/10 PLC patients ($p<0.10$, Chi-square). All four patients with serositis, all randomized to PLC, failed to respond. An imbalance at baseline ($p<0.05$) was noted in the physician's global, with the DHEA group scoring worse.

Secondary analyses showed the following:

	DHEA	PLC	p-value*
	mean (sem)		
SLEDAI	15.8 (7.7)	9.4 (6.0)	0.06
SLAM	15.3 (4.8)	11.3 (4.4)	0.08
Investigator global	66.1 (23.0)	44.4 (19.0)	0.04
Patient global	66.8 (29.6)	52.6 (23.)	0.26
Prednisone dosage	49.9 (18.3)	44.0 (22.7)	0.58
Pts. on I'suppres. rx	5 (50%)	6 (60%)	
ESR	83.6 (38.4)	68.8 (39.5)	0.42
Anti-DNA**	262 (300)	541 (106.5)	0.46

* by 2-tailed Student t-test

** reciprocal of titer by Crithidia method

As pointed out in the publication of this trial, the SLEDAI weighs nephritis more than serositis, so the SLEDAI result could be from the greater number of nephritis patients in the DHEA arm at baseline.

Adverse events included the following:

	DHEA	PLC
Acne	6	3
Mood change	1	3
Menses change	8	3
Hirsutism	4	2
Headache	4	4
Insomnia	2	2
Decreased libido	1	1
Hair loss	0	1

3. Non-IND Study GBL96-01 (Taiwan)

This non-IND study was added to the evidence prepared for the Arthritis Advisory Committee. Additional documentation would be needed for the agency to perform a full review of the study results. Study GBL96-01 was a six month study in women with mild to moderate lupus (up to 10mg/day of prednisone, and a SLAM score of at least seven, later amended to also require a SLEDAI of more than two). 120 patients were randomized to DHEA 200mg/day or PLC. The primary endpoint was the mean change in the SLAM at six months. There were a number of secondary endpoints.

The two arms were well balanced at baseline. Three of 61 DHEA patients and four of 59 PLC patients withdrew before completion. The mean (medium) change in the SLAM for the DHEA arm was 2.6 (2.5), compared to 2.0 (2.0) for the PLC arm, p=0.355. Other variables – SLEDAI, patient VAS, and physician VAS – showed a significant p value for only the patient VAS (p=0.005). Percent responders (analogous to the endpoint used in GL95-02) was 20/59 for DHEA, compared to 18/57 for PLC, p=0.792. A time to flare analysis was significant at 0.044.

4. Open label study (van Vollenhoven #3): Use of DHEA 200mg/day in female patients with SLE (J Rheum 25:285-9, 1998)

This was an open label study of 50 females, 37 pre- and 13 post-menopausal, with SLE treated with 50-200mg/day DHEA conducted at Stanford University by Dr. van Vollenhoven. Thirty-four patients completed six months therapy, and 21 patients completed 12 months. A reduction in disease activity to a mean of 2 on the SLEDAI scale compared to baseline mean of 8 for patients-on-treatment was seen. Patient and physician global assessments also improved, and the average

prednisone dose was reduced from 12mg to 5mg/day. No effect was seen on laboratory parameters. Side effects seen included acne 54%, hirsutism 18%, alopecia 10%, perspiration 8%, oily skin 4%, and greasy hair 2%; menstrual irregularity 4%, breast tenderness 6%; headache 4%, vertigo 4%, mood alteration 6%; and weight gain 6%, striae 2%, worsening bronchospasm 2%.

5. **Open label PK/clinical study (Van Vollenhoven #4): Escalating doses from 50mg/day to 600mg/day of DHEA in females with SLE (J Rheum 25:2352-6, 1998)**

This study also was conducted at Stanford University and attempted to explore relations between blood hormonal levels induced by DHEA therapy and efficacy and toxicity. Twenty-three women with mild SLE, without renal or CNS involvement, and requiring no more than 10mg/day of prednisone, were treated with monthly escalating dosages of DHEA for 6 months (50mg to 600mg/day), provided no side effects occurred and disease activity ("the absence of remission" defined as a SLAM<4) remained. Patients were seen monthly and assessed as follows: SLEDAI, SLAM, HAQ, KFSS, physician and patient global scales, VAS pain, DHEA, DHEA-S, and serum testosterone levels.

Maximal dose achieved was 100mg/d in 4, 200mg/d in 4, 400mg/d in 11, and 600mg/d in 3 patients. Escalation was stopped for the following reasons:

Max. Dose DHEA	No. patients	Side effects	Remission
100mg/day	4	3	1
200mg/day	4	3	1
400mg/day	11	6	5
600mg/day	3	0	3

A trend toward dose linearity was seen for dose versus serum DHEA (r = 0.531 (0.386, 0.651)) and for dose versus serum DHEA-S (r = 0.608 (0.478, 0.712)). Weak and negative pharmacodynamic correlations were seen between DHEA levels and SLAM (r=-0.290 (-0.45, 0.11)), and between DHEA-S levels and SLAM (r= - 0.16 (-0.34, 0.03)). In view of the possibility of an "optimum" curve, with declining efficacy at supra-optimal serum levels, a second order linear regression analysis was done of DHEA-S versus SLAM. There was seen in this second order model the suggestion of an optimal serum DHEA-S level of about 1000ugm/dl (r=0.25, p<0.04). There did not appear to be a pharmacodynamic effect for toxicity, looking at associations between DHEA, DHEA-S, or testosterone and acne, the most common adverse event seen, but there clearly may have been inadequate power to detect such effects.

Kent Johnson, M.D.
April 13, 2001